CpE 645 Introduction to Image Processing and Coding

Prof. Hong Man

Department of Electrical and Computer Engineering
Stevens Institute of Technology

STEVENS Institute of Technology

Visual Information Environment Laboratory

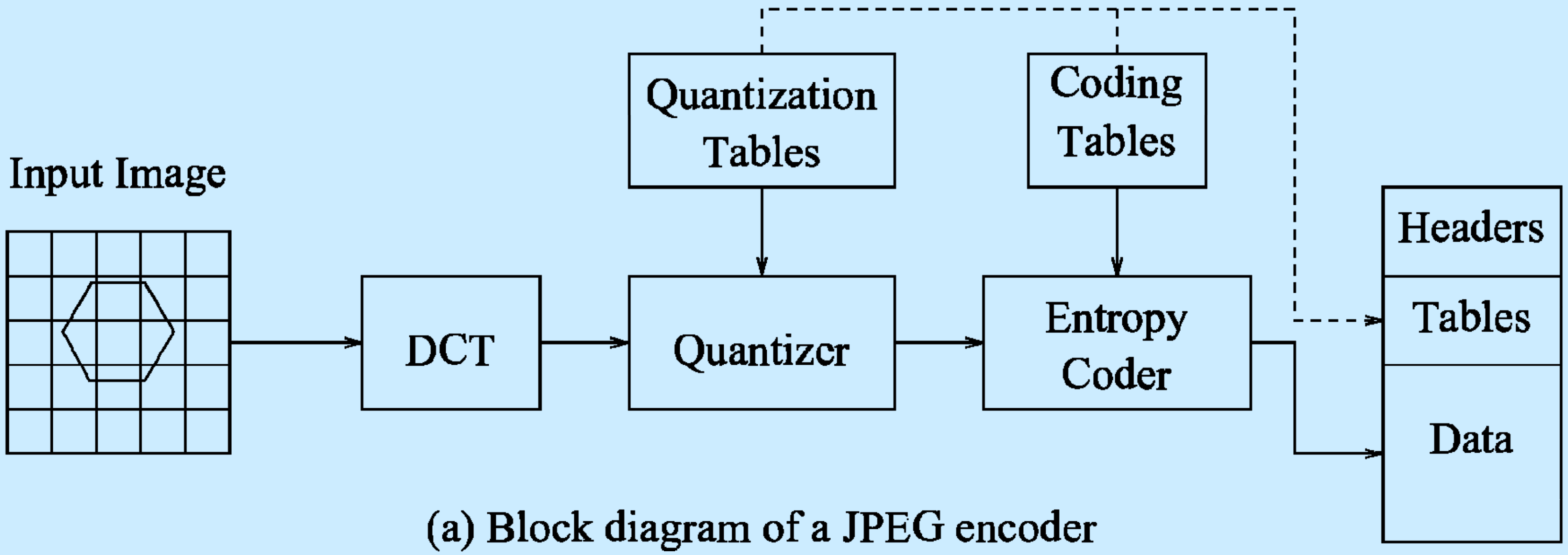
Lossy Still Image Coding

- DPCM approach: earliest and simplest non-PCM method.
- Transformation-Quantization-Entropy Coding approach:
  - Transform stage:
    - DCT based: JPEG
    - Subband/wavelet based: JPEG-2000
  - Quantization stage:
    - Scalar Quantization,
    - Vector Quantization.
  - Entropy coding stage:
    - Huffman coding,
    - Arithmetic coding.

The JPEG Standard

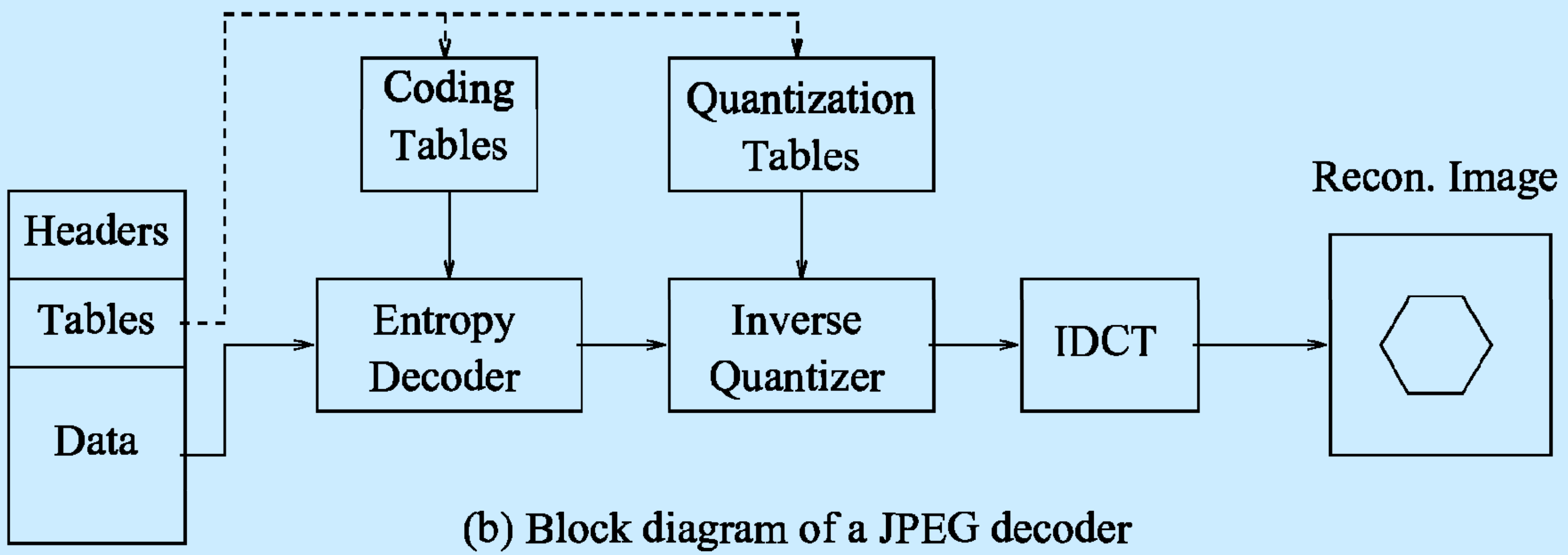
- JPEG: Joint Photographic Experts Group (ISO10918, ITU-T T.81).
- Requirements:
  - Generic still image compression
  - Modest to low software/hardware complexity,
  - Sequential, Progressive and layered coding.
- Features:
  - Psychovisual-based quantization,
  - Sequential, progressive/ hierarchical coding modes,
  - Relatively low memory requirement.

The JPEG Standard (*cont.*)



(a) Block diagram of a JPEG encoder

The JPEG Standard (*cont.*)



(b) Block diagram of a JPEG decoder

The JPEG Standard (*cont.*)

- Input of the coding process: 8x8 block, 8 bits/pixel.
- *Example:*

$$\begin{bmatrix} 175 & 172 & 172 & 173 & 169 & 174 & 171 & 167 \\ 171 & 170 & 170 & 174 & 177 & 175 & 171 & 166 \\ 167 & 168 & 172 & 174 & 169 & 172 & 174 & 161 \\ 149 & 149 & 156 & 153 & 148 & 149 & 155 & 141 \\ 101 & 106 & 105 & 107 & 99 & 92 & 106 & 97 \\ 73 & 74 & 72 & 71 & 65 & 70 & 79 & 72 \\ 72 & 71 & 67 & 62 & 62 & 72 & 80 & 71 \\ 68 & 70 & 64 & 58 & 62 & 73 & 74 & 67 \end{bmatrix}$$

The JPEG Standard (*cont.*)

- 8x8 DCT: the most computational demanding stage.
- *Example:*

$$\begin{bmatrix} 31144 & 52 & 35 & 86 & -294 & 264 & -137 & 14 \\ 8039 & 100 & -258 & 11 & 108 & 10 & 38 & -12 \\ -387 & -89 & 117 & 156 & -12 & -103 & 90 & -52 \\ -1778 & 47 & 96 & -78 & 59 & 10 & -31 & -19 \\ 164 & 80 & -7 & -43 & -8 & 0 & -6 & 12 \\ 551 & -17 & 54 & 18 & -67 & 64 & 47 & -3 \\ -152 & 4 & 6 & -16 & -30 & 24 & 15 & -34 \\ -83 & 19 & 12 & -37 & 6 & 15 & -46 & -6 \end{bmatrix}$$

The JPEG Standard (*cont.*)

- Quantization is performed using a quantization table (the Q-table).
- *Example:*

$$\begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

The JPEG Standard (*cont.*)

- Let the **8x8** DCT coefficients $X[k_1, k_2]$ be the input of the quantization, and $Q[k_1, k_2]$ be the Q-table, the weighted DCT coefficients becomes

$$Y[k_1, k_2] = X[k_1, k_2]/Q[k_1,k_2].$$

- The output of the quantization stage is

$$\hat{Y}[k_1, k_2] = \mathrm{round}(Y[k_1, k_2]).$$

- A Q-table is designed to minimize the mean square error or maximize the visual quality.
- For a specific coding process, the prototype Q-table can be scaled with a factor in (**1~100**) to achieve different quality and compression ratios.

The JPEG Standard (*cont.*)

- The DCT coefficients are weighted by the Q-table.
- *Example:*

$$\begin{bmatrix}
1946.5 & 4.7 & 3.5 & 5.4 & -12.3 & 6.6 & -2.7 & 0.2 \\
669.9 & 8.3 & -18.1 & 0.6 & 4.2 & 0.2 & 0.6 & -0.2 \\
-27.7 & -6.9 & 7.3 & 6.5 & -0.3 & -1.8 & 1.3 & -0.9 \\
-127.0 & 2.8 & 4.4 & -2.7 & 1.2 & 0.1 & -0.4 & -0.3 \\
9.1 & 3.6 & -0.2 & -0.8 & -0.1 & 0 & -0.1 & 0.2 \\
23.0 & -0.5 & 1.0 & 0.3 & -0.8 & 0.6 & 0.4 & 0 \\
-3.1 & 0.1 & 0.1 & -0.2 & -0.3 & 0.2 & 0.1 & -0.3 \\
-1.2 & 0.2 & 0.1 & -0.4 & 0.1 & 0.1 & -0.4 & -0.1
\end{bmatrix}$$

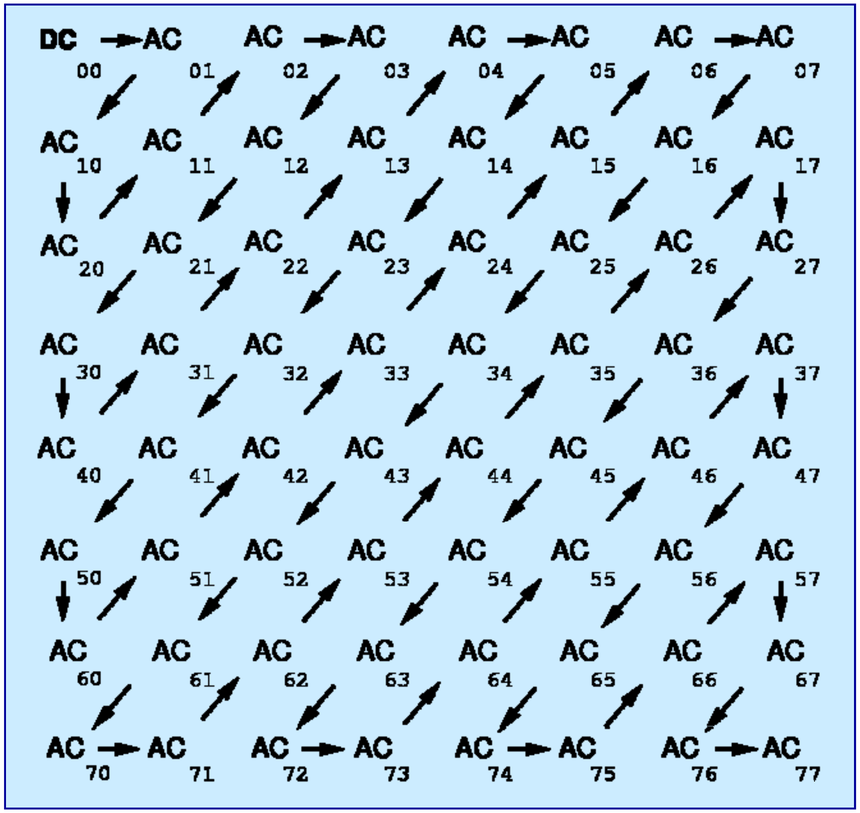
The JPEG Standard (*cont.*)

- The weight DCT coefficients are rounded to integers. The resulting block is then coded through the entropy coding.
- *Example:*

$$
\begin{bmatrix}
1947 & 5 & 4 & 5 & -12 & 7 & -3 & 0 \\
670 & 8 & -18 & 1 & 4 & 0 & 1 & 0 \\
-28 & -7 & 7 & 7 & 0 & -2 & 1 & -1 \\
-127 & 3 & 4 & -3 & 1 & 0 & 0 & 0 \\
9 & 4 & 0 & -1 & 0 & 0 & 0 & 0 \\
23 & 0 & 1 & 0 & -1 & 1 & 0 & 0 \\
-3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

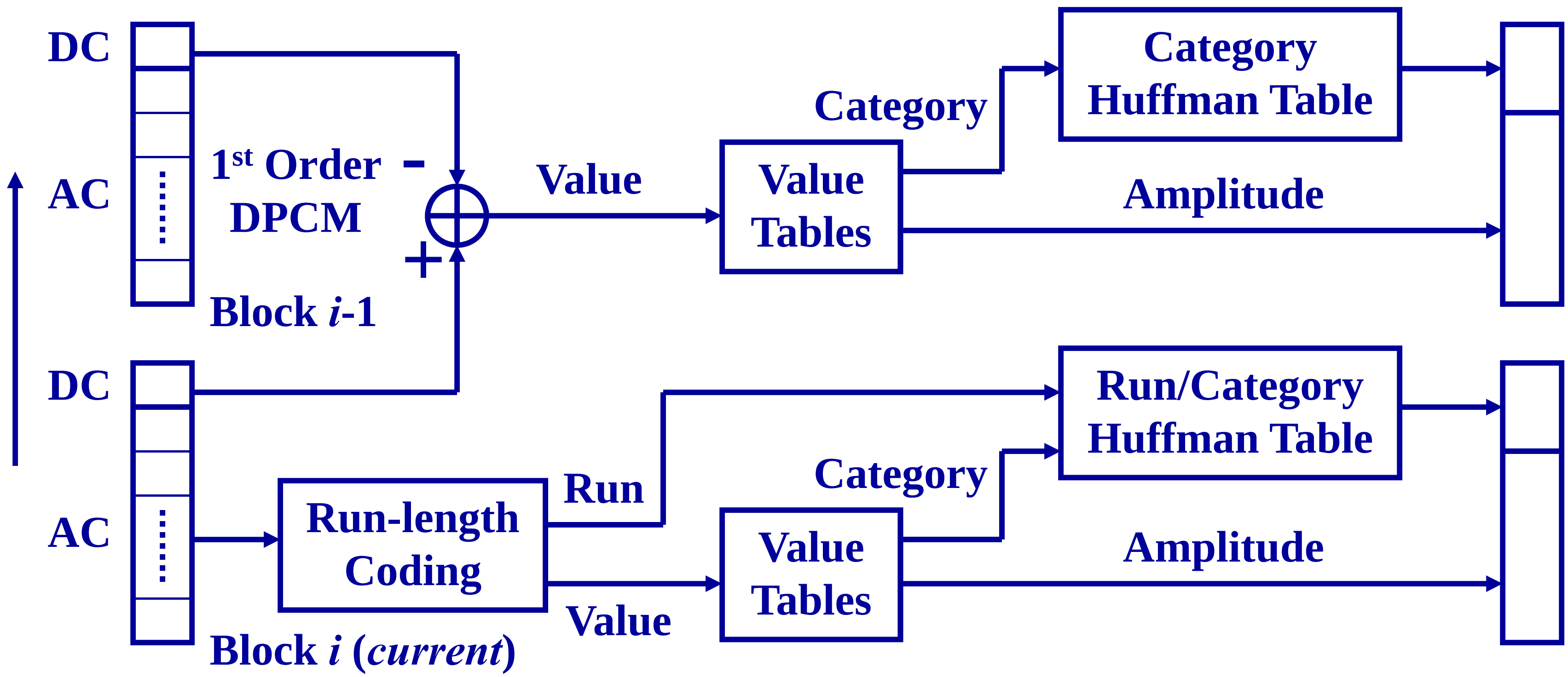
The JPEG Standard (*cont.*)

- Each quantized block is scanned through a zig-zag pattern and converted into 1-D data streams.



The JPEG Standard (*cont.*)

- The DC and AC bit streams are coded through Huffman coding together with DPCM and run-length coding.



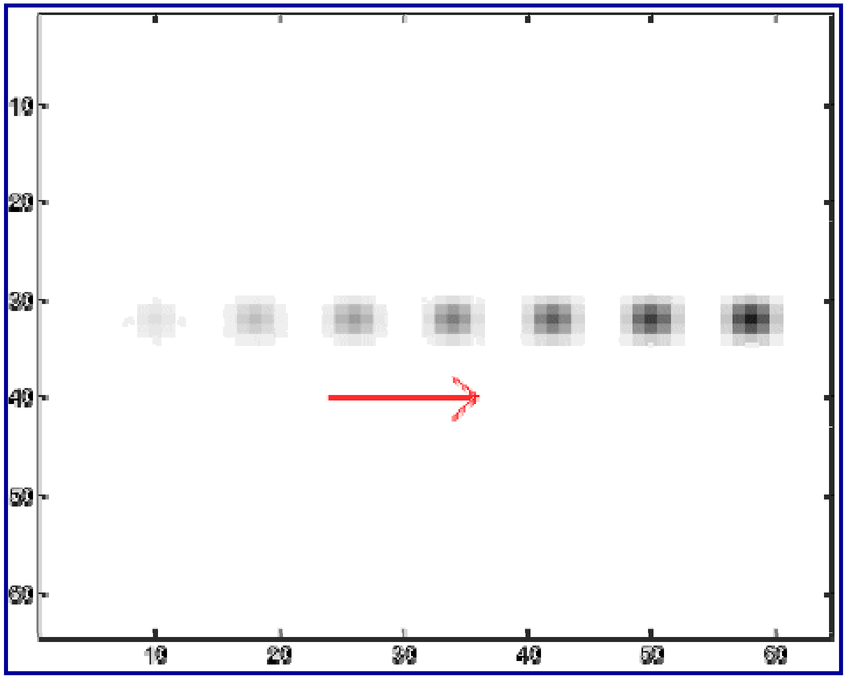
Digital Video

- Digital video is a sequence of still images (or frames) displayed at rapid succession.
- Digital video usually presents huge data volume:
  - NTSC video (720x480, 3-color, 30 fps): 249 Mbps.
  - HDTV video (1920x1080, 3-color, 30 fps): 1.49 Gbps.
  - Super-35 video (4096x3112, 3-color, 30 fps): 9.18 Gbps.
- Digital video compression is to exploit the data redundancies in both **spatial** domain (*i.e.* image compression) and **temporal** domain.

Example of Video Frames

Motion Representation

- Motion of an object or scene can be simplified and modeled as some ***linear motion***, in which a motion vector can be used to represent the motion.



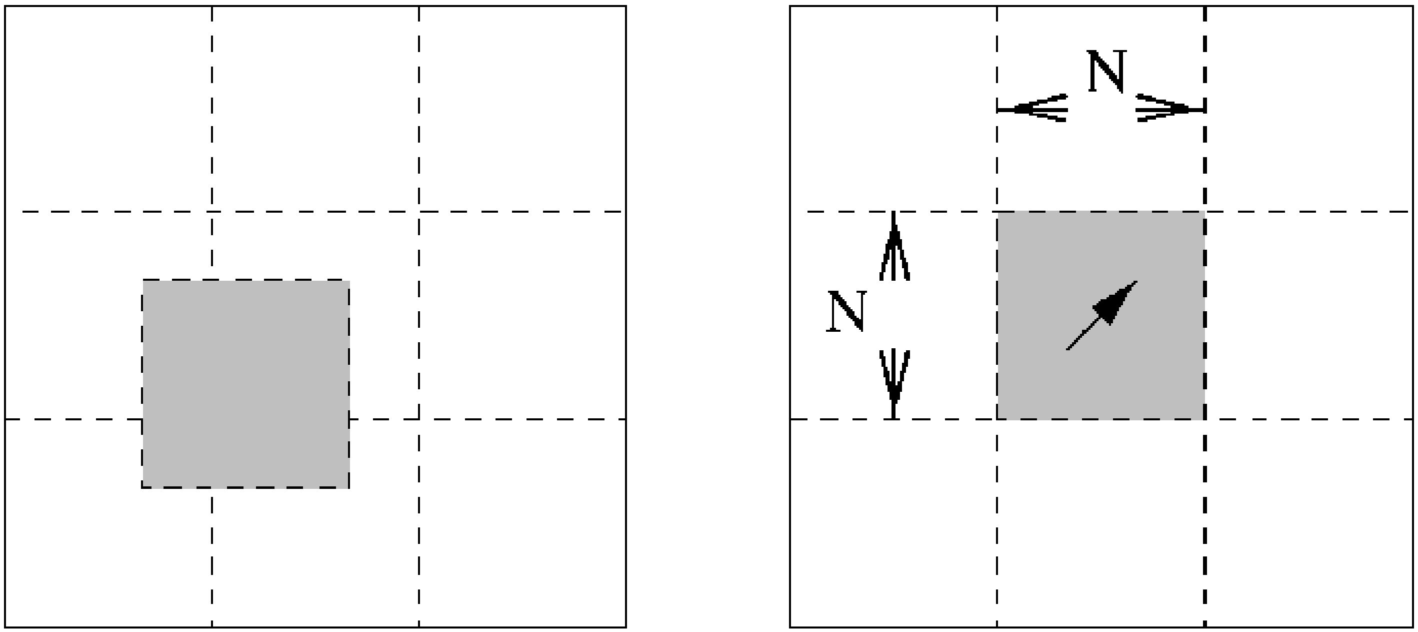
Block Motion Estimation

- In a complex nature scene, it is difficult to track all the moving objects.
- Block based motion estimation divides each video frame into fixed-sized non-overlapping blocks, and treat each block as an object to track its motion.
- A block in **current** frame is compared with all possible regions in the **reference** frame, this search will find a best match and produce a **motion vector** representing the displacement between the current block and its best match.
- The reference frame can be previous frame or future frame or both.

Block Motion Estimation (*cont.*)

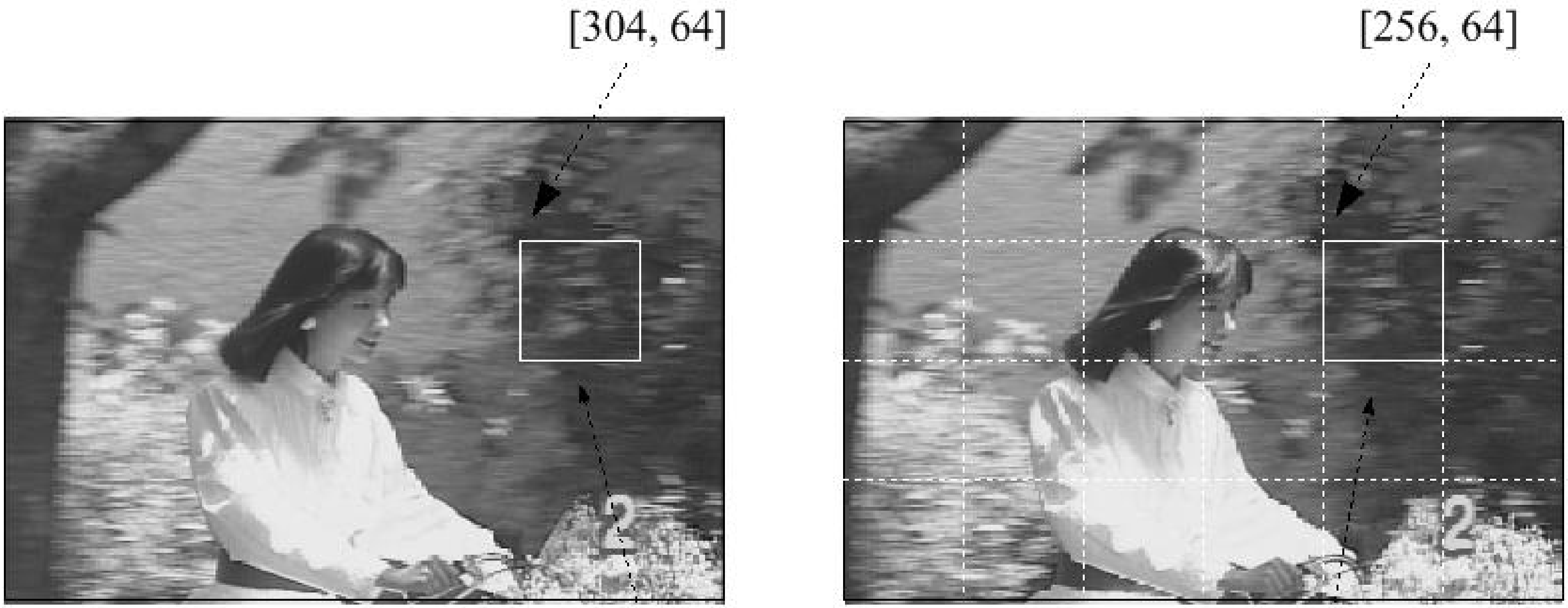
previous frame
frame i-1

current frame
frame i



(b) (a)

Block Motion Estimation (*cont.*)



reference frame

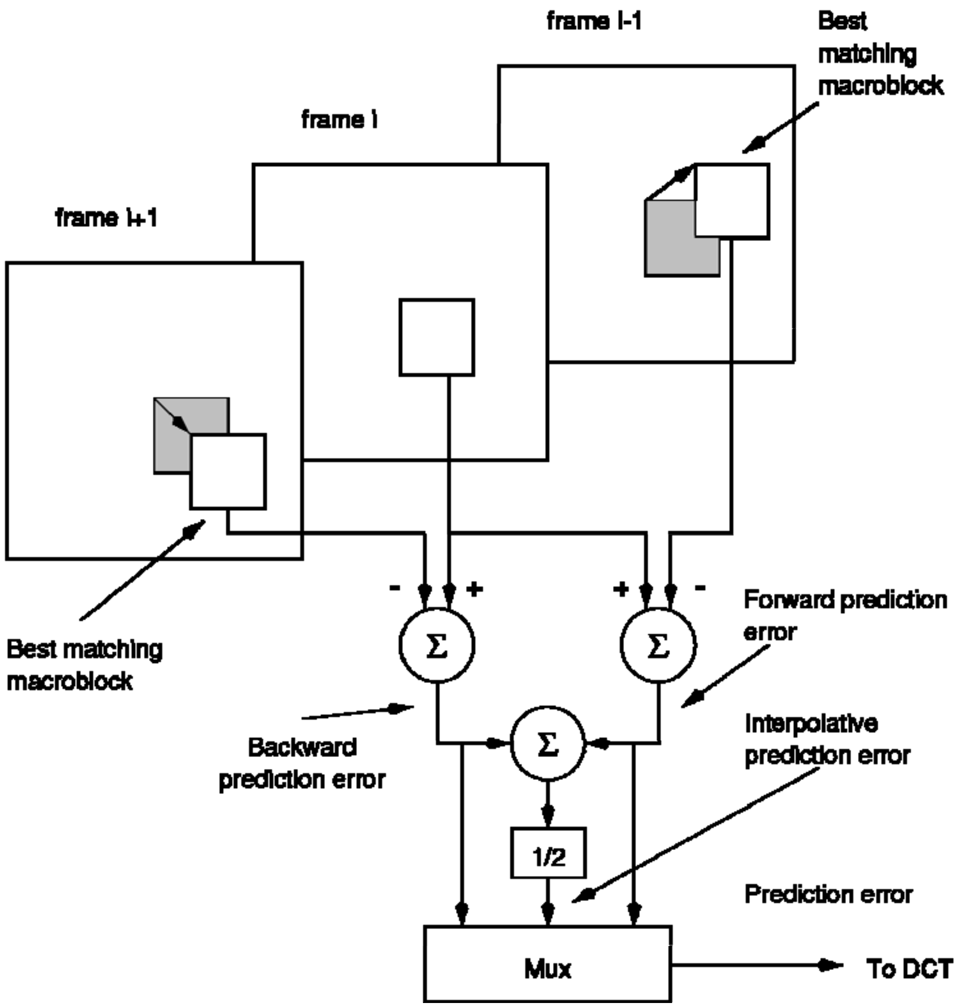
current frame

this reference block is selected as the
prediction block for this current block
with a motion vector of [48, 0]

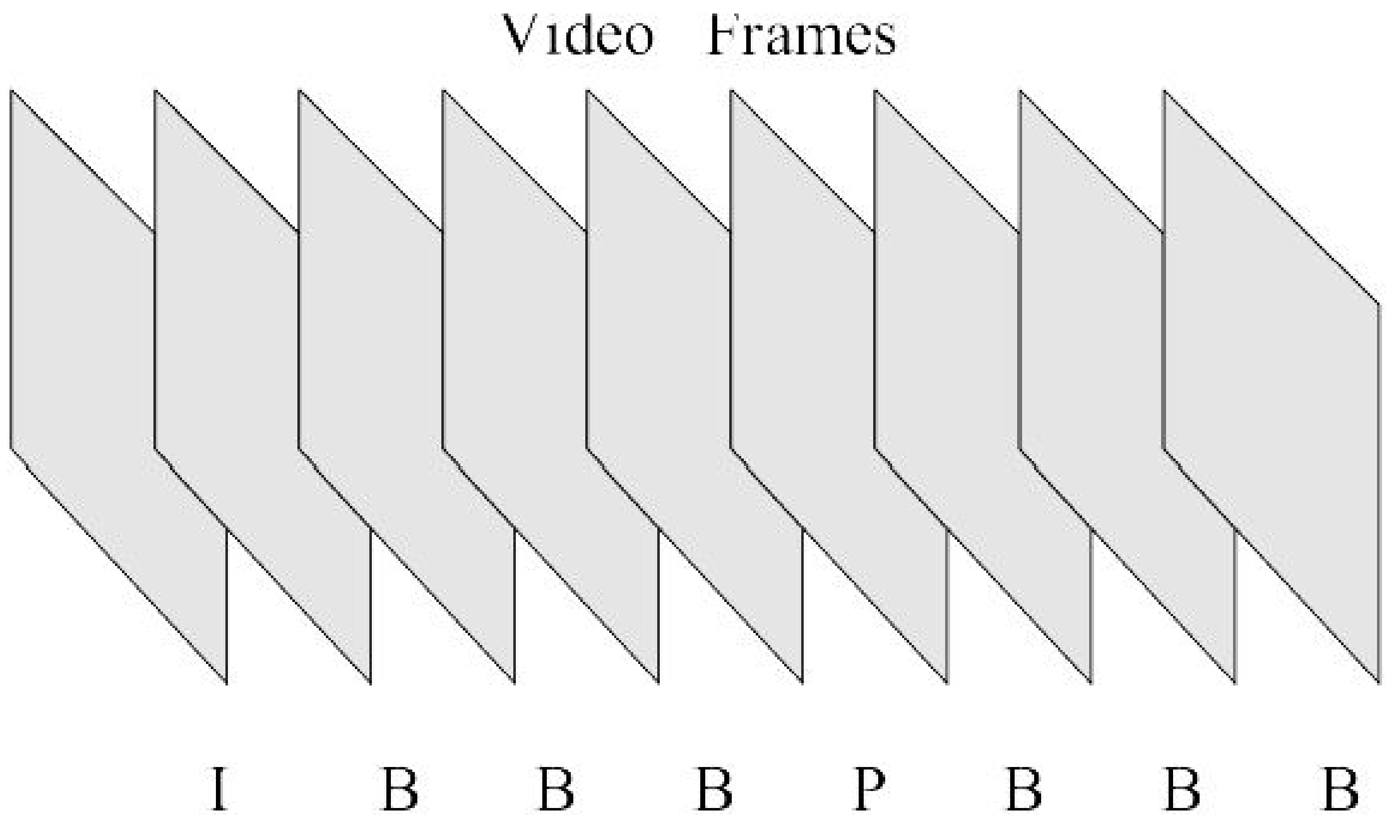
Motion Compensated Prediction

- **Motion compensated prediction** is a special prediction method based on the knowledge of motion in a video sequence. It can be used in a prediction coding structure.
- The procedure:
  - Block-based motion estimation, generating a motion vector for each block.
  - Moving the best match of each block to its current position to form a prediction block.
  - Deferential coding of each frame based on the prediction frame formed by all the prediction blocks.
  - Sending all the motion vectors to the decoder to generate the same prediction frame for decoding.

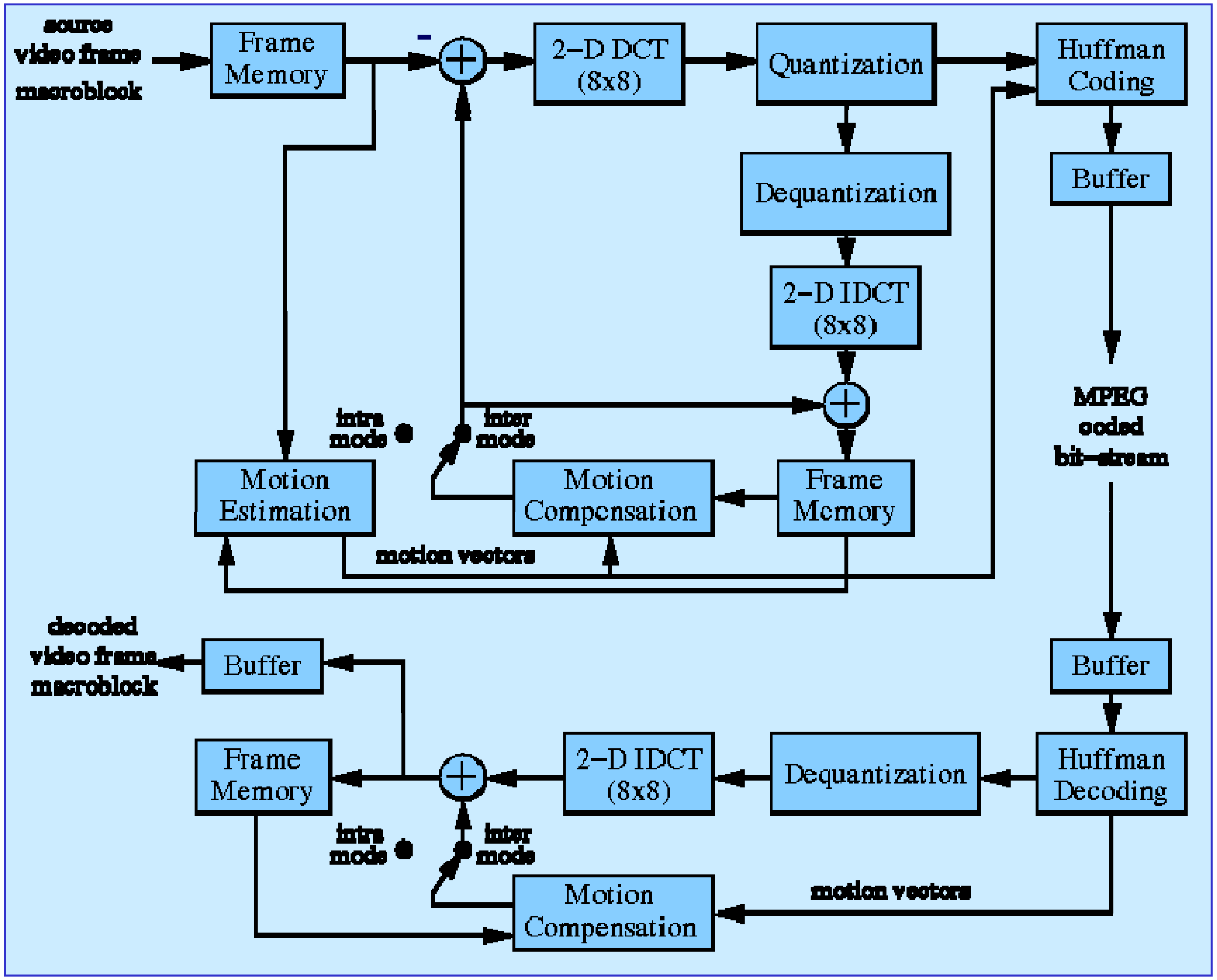
Bi-directional MC Prediction



I, P, and B Frames



Moving Pictures Experts Group (MPEG)



STEVENS Institute of Technology

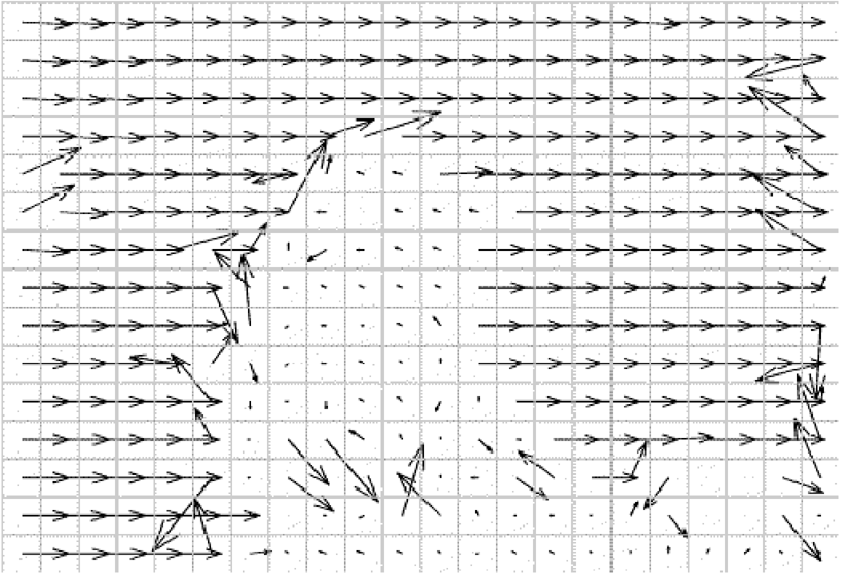
Visual Information Environment Laboratory

MC-Prediction: Example

Previous frame

Current frame

MC-Prediction: Example (*cont.*)



Motion vector field

Prediction of current frame

MC-Prediction: Example (*cont.*)

Residual frame with MC

Residual frame with direct difference

ITU-T H.261

- ITU-T H.261—The earliest video coding standard (1990).
- It is targeted for video teleconferencing over ISDN lines.
- Input formats: CIF (288x352) and QCIF (144x176).
- Bit Rate: multiples (1~30) of 64 kbits/sec.
- System delay is critical.
- Forward prediction only: I-, P-frames.
- A loop filter is used to reduce block artifacts.
- Bit stream structures: frames, group of blocks, macroblocks and block(8x8).

MPEG-1

- MPEG-1 — ISO standard 11172
  (11172-1: system, 11172-2: video, 11172-3: audio).
- Coding of moving pictures and associated audio for digital storage media applications.
- MPEG-1 is based on ITU-T H.261.
- It is optimized for 30 fps SIF format (352x240).
- Bit rate: ~1.5 Mbits/sec.
- System delay is not critical.
- Bi-directional motion estimation: I-, P-, B-frames.
- D-frame (DC) is supported for fast frame access.
- Bit stream structure: sequence, group of pictures (GOP), picture, slice, macroblock and block(8x8).

MPEG-2

- MPEG-2 — ISO standard 13818
  (13818-1: system, 13818-2: video, 13818-3: audio).
- Generic coding of moving pictures and associated audio.
- MPEG-2 is backward compatible to MPEG-1.
- Five **profiles** (algorithms), four **levels** (parameter sets) are defined for selecting in any specific application:
  - Profiles: simple, main, SNR-scalable, spatially-scalable, high.
  - Levels: low (352x288, 30fps, 4Mbps), main (≤720x576, 30fps, 15Mbps), high-1440 (≤1440x1352, 60fps, 15Mbps), high (≤ 1920x1152, 60fps, 80Mbps).

ITU-T H.263

- ITU-T H.263: Superior performance for very low bit rates.
- It is targeted for video telephony.
- H.263 is closely related with H.261.
- Bit Rate: below 64 kbits/sec.
- Forward prediction only: I-, P-frames.
- Hierarchical motion estimation.
- Differential coding of motion vectors.
- Support overlapping of motion vector blocks.

MPEG-4

- MPEG-4 defines a set of coding tools and a syntactic description for audio-visual objects (AVOs).
- Each of the AVO can be coded independently.
- Coding of object shapes is essential.
- Object-based motion compensation.
- Panoramic still background is coded using **sprite** coding.
- Still texture image is coded using wavelet coding.
- Supporting SNR-, spatial- and object-scalability.
- Server-side interaction: content-based manipulations.
- Client-side interaction: AVO modifications.
- Much more…

DVD-Video

- DVD specification is defined and maintained by DVD Forum (http://www.dvdforum.org/forum.shtml).
- DVD-Video storage capacity 17 Gbyte if two layers on both sides of the disk are utilized.
- Specifications and features
  - 133 minutes of high quality MPEG-2 encoded video with multi-channel Dolby Surround AC-3 audio can be stored on one layer on one side (around 4 Gbyte).
  - Support widescreen, letter box and pan & scan video formats. (4:3 and 16:9 aspect ratios).
  - Up to 8 tracks of digital audio for multiple languages, each with as many as 8 channels.
  - Up to 32 subtitle/karaoke tracks.

DVD-Video

- Specifications and features
  - Menus and program chains for user interactivity
  - Up to 9 camera angles to give the user more choice
  - Digital and analogue copy protection
  - Parental control for protection of children
  - Special effects playback: freeze, step, slow, fast, and scan (no reverse play or reverse step).
  - Random play and repeat play.
  - Programmability (playback of selected sections in a desired sequence).

DVD-Video

- Content Scrambling System (CSS)
  - CSS is a data encryption and authentication scheme intended to prevent copying video files directly from DVD-Video discs
  - The CSS decryption algorithm exchanges keys with the drive unit to generate an encryption key that is then used to obfuscate the exchange of **disc keys** and **title keys** that are needed to decrypt data from the disc.
  - DVD players have CSS circuitry that decrypts the data before it's decoded and displayed.
  - In October 1999, the CSS algorithm was cracked and posted on the Internet, triggering endless controversies and legal battles

DVD-Video

- DVD regional codes:
  - Motion picture studios want to control the home release of movies in different countries because theater releases aren't simultaneous
- More info about DVD-Video

http://www.disctronics.co.uk/technology/dvdvideo/dvdvid_intro.htm

http://www.dvddemystified.com/dvdfaq.html